National Security Agency/
Central Security Service

INFORMATION
ASSURANCE
DIRETORATE

Defending Risky Electronic Access Points into a “Closed” Industrial Control System (ICS) Network Perimeter

This publication is the third in a series intended to help Industrial Control System (ICS) owners and operators in need of improving the security posture of their systems. This document will focus the reader on aspects of system security within a “closed” ICS network perimeter, and give them a systematic approach for implementing the access control concept of Least Privilege. By restricting the accesses for process automation devices to and from only authorized subjects, ICS security can be built in a layered manner. Starting at the heart of a closed system and controlling access to devices in the system network out to devices which straddle the physically-controlled system network boundary, ICS owner’s can build hardened “closed” ICSs which both logically and physically address the perceived security threats.

Control Systems Division

Contents

1. Introduction

Industrial control systems (ICS) control and monitor complex industrial processes like petroleum refinement, chemical production, product manufacturing, and electric power generation and transmission. Much of the global critical infrastructure is dependent on industries that employ networked ICS systems. Sabotage or disruption of these industries can have wide-ranging negative effects including loss of life, economic damage, property destruction, or environmental pollution. Our reliance on ICS networks, coupled with the potential for widespread damage makes them attractive targets for electronic attack.

Electronic attacks are most likely to enter protected networks through communications links that are significantly exposed to unauthorized outsiders. Electronic access points that create direct or indirect connections between critical ICS network segments and publicly-accessible network infrastructures like the Internet or public switched telephone network (PSTN) are of particular concern. In this publication, we present network architectures and cryptographic security technologies that can be employed to block unauthorized network traffic from entering protected ICS network segments through these high-risk network connections.

2. Identifying Risky Network Connections

The goal of a network security improvement effort is to limit access to electronically-accessible functions, actions, and sensitive data to those devices or users that either have privilege to access the action/data or are trusted not to access the action/data. Central to this concept is the ability to profile all users and devices on the network and to identify the rights and privileges that must be granted to them. In this publication, the term “insiders” refers to users for which some authorized role, or collection of rights and privileges, can be identified. Insiders have defined, authorized roles on the protected network. In contrast, those potentially unidentified users that do not have any rights or privileges on the protected network are referred to as “outsiders”.

A closed network is any network that is “significantly” protected from physical and electronic access by unauthorized outsiders. The following qualities minimize the likelihood of unauthorized physical or electronic intrusion by outsiders and, thus form the definition of a closed network:

- Assets on the closed network operate within an uninterrupted, defended physical perimeter (locked or monitored building complex, unbroken fence line with locked gates, etc.)
- Assets on the closed network are electronically connected to one another via continuous or intermittent, wired communications links carried on media located entirely within the secure physical perimeter (networked wireless devices do not fit this definition)
- All users that can access networked assets on the closed network have known, definable trust and privilege profiles

Electronic connections into, or out of a network segment that violate the above definition are very often the riskiest links on an ICS network due to the risk of outside intrusion. For example, wide area network (WAN) connections that leave the secured physical perimeter may be susceptible to unauthorized intrusion by outsiders with access to the WAN media or infrastructure. Such intrusions often include the ability for an attacker to intercept sensitive data from the compromised network link and/or transmit malicious attack data to the assets serviced by the link.

Network links with significant exposure to electronic attack and significant potential for damage exhibit the highest threat potential. Of particular concern are direct or indirect connections between networked, critical ICS assets and globally-accessible communications infrastructures like the Internet or dialup Public Switched Telephone Network (PSTN). These connections often expose ICS communications services that allow the modification of critical ICS functions, often including the ability to reprogram critical ICS devices or operate potentially damaging control points (e.g. pump controls, switches, valves, or breakers).

In this publication, we will discuss techniques for creating a strong network perimeter to reduce the risk of undetected, unauthorized access by outsiders. Risky network connections within an entirely closed network may also benefit from the security recommendations discussed in this publication. For example, electronic connections that link ICS network segments that exhibit dramatically different user privilege profiles or security requirements can be secured and isolated using the techniques discussed below.

3. Techniques and Technologies for Defending High-Risk Network Links

In this section, we will introduce some common technologies and practices that can be used to secure high-risk network links.

3.1. Create Demilitarized Zone (DMZ) Subnetworks to Isolate High-Risk Internet Protocol (IP) Network Connections

It is very common in modern network design to have to share communications services between two IP network segments that exhibit dramatically dissimilar purpose or security requirements. For example, most corporate IP networks must connect to the Internet to allow services like e-mail and website access to flow between its users and the Internet. The Internet is a largely unregulated environment rife with potential attackers and malicious network traffic, while a typical corporate network contains workstations and servers storing sensitive data and providing services essential to the daily operation of the business. Any connection to the Internet infrastructure can expose protected network assets to targeted or random Internet-borne electronic attack and, therefore, must be properly secured and monitored. Fortunately, effective network design principles and techniques exist to minimize the possibility of attack propagation between connected network segments.

One very effective network architecture employs an intermediate LAN segment, called a Demilitarized Zone (DMZ) to isolate and secure a risky connection between two LAN segments. Figure 1 shows two connected LAN segments (labeled “LAN 1” and “LAN 2” in the figure) separated by an intermediate DMZ security zone (labeled “DMZ LAN” in the figure). The connections between the two LAN segments and the DMZ are protected by two independent traffic filtering firewalls[1] (represented as brick walls in the figure).

[1] It is also acceptable to use a single, three-interface firewall. There are, however, some minor security benefits associated with using two, two-interface firewalls from different manufacturers. A vulnerability present in one of the

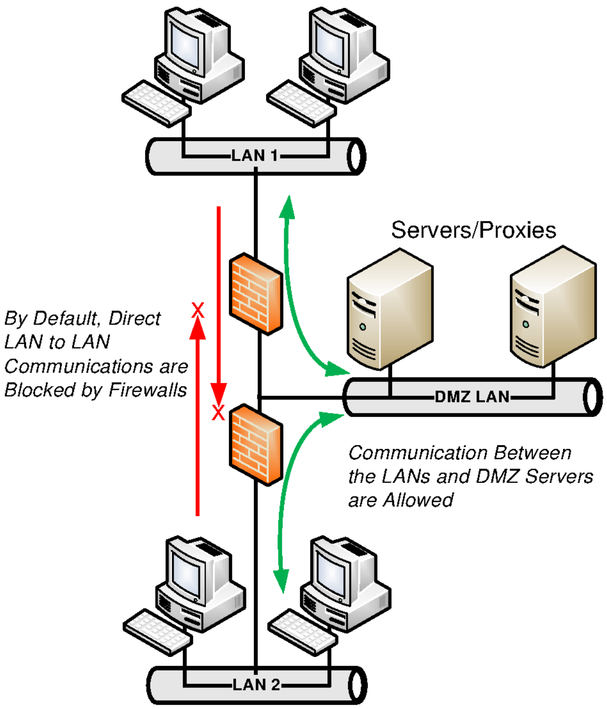

LAN 1
Servers/Proxies
By Default, Direct LAN to LAN Communications are Blocked by Firewalls
X
X
DMZ LAN
Communication Between the LANs and DMZ Servers are Allowed
LAN 2

Figure 1 A Connection Between Two LANs can be Effectively Controlled and Secured by Creating an Intermediate Demilitarized Zone (DMZ) LAN Segment, Protected by Dual Firewalls

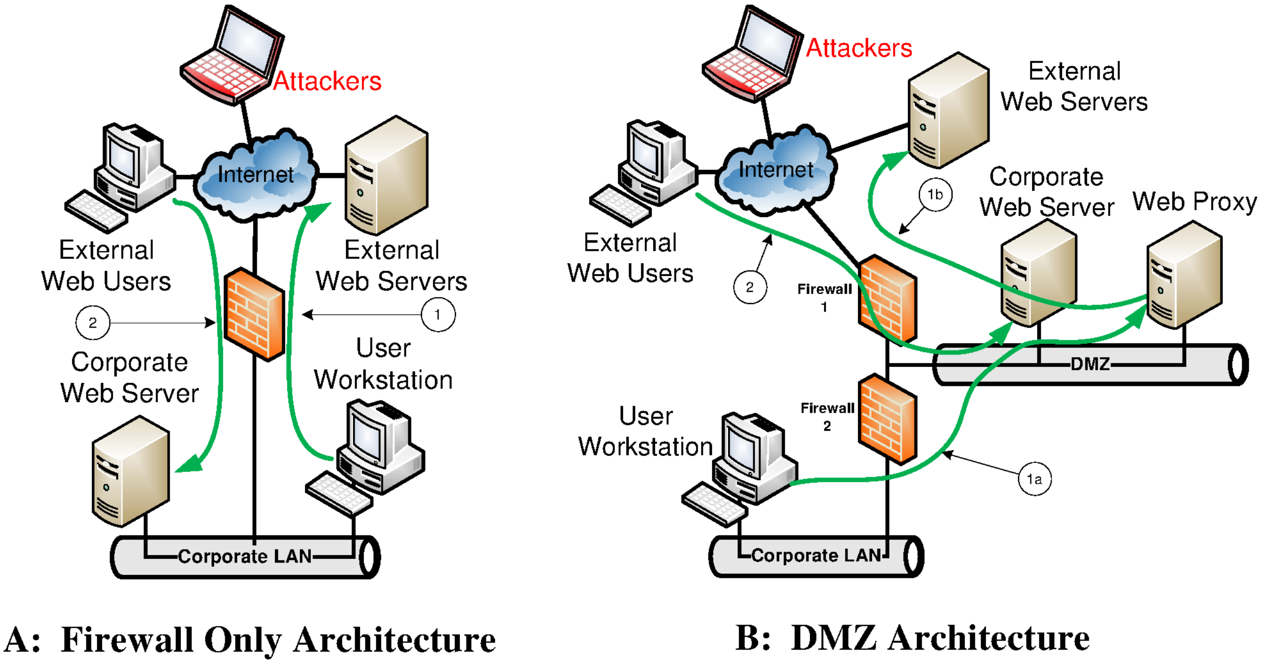
Firewalls are network security products that regulate the Internet Protocol (IP) network traffic flowing through them. Configurable rule sets in the firewall specify which types of traffic are allowed to be routed between its network interfaces and which types are to be dropped/ignored. By default, a firewall's rule set should block all traffic from flowing through it. Network security personnel then configure the firewall by entering exception rules that allow specific traffic through the device. Exception rules can allow traffic based on its IP source and destination addresses, TCP source and destination ports, and other specific characteristics. For example, an exception rule can be entered that allows only incoming Telnet traffic (TCP port 23) to flow between a specific host on one port (e.g. IP address 192.168.1.1), and another specific host (e.g. IP address 192.168.5.5) on its other port: in the absence of any other exception rules, all other traffic received by the two firewall ports will be dropped/ignored. Stateful firewall models can even differentiate from which port the connection originated, and support exception rules that take the directionality of the TCP/IP connection (which party initiated the session) into account. For example, the Telnet exception rule described above can be modified on a stateful firewall to only allow Telnet connections to be initiated from one port, eliminating the possibility of unsolicited Telnet session requests from the other.

Traffic filtering rules in the firewalls in Figure 1, effectively enforce maximum isolation between "LAN 1" and "LAN 2" while still allowing data to safely flow between the LAN segments and the servers on the DMZ. The intermediate DMZ provides a security zone from which servers/proxies provide the required communications services without requiring direct communications between "LAN 1" and "LAN

firewalls is less likely to also be present in the other firewall, thus preserving some security in the event of a severe product vulnerability. In addition, independent firewalls can be managed from separate network segments, eliminating the chance of compromising the firewall administration interfaces from a single location.

2" (e.g. a TCP/IP session directly between a host on "LAN 1" and a host on "LAN 2"). The firewalls are configured such that only the most essential communications services are allowed to flow between the LAN segments and the DMZ.

The best way to explain the behavior and potential security benefits of the DMZ architecture is to use a common corporate Information Technology (IT) problem as an example. In Figure 1, we have connected two LAN segments via an intermediate DMZ. In general, users on the "LAN 2" segment must be able to retrieve information from "LAN 1". In the following example, we will use external Internet web access to represent this need: users on the corporate network must be able to access external web servers on the Internet in order to browse websites. Likewise, users on "LAN 1" must be able to retrieve information from "LAN 2". In our example, we will use a corporate website server to represent this need: users on the Internet must be able to access a web server somewhere on the protected corporate network infrastructure in order to view the company website. Figure 2 shows the two competing architectures that we will analyze for our example: A) shows a less effective, two-interface firewall architecture, and B) shows a more secure architecture employing an intermediate DMZ segment protected by two independent, two-interface firewalls.



A: Firewall Only Architecture

B: DMZ Architecture

Figure 2 Competing Configurations Showing the Benefits of a DMZ Architecture: A) A Single, Two-Interface Firewall Architecture, and B) A DMZ Architecture

We will begin by analyzing the single firewall architecture shown in Figure 2A (the left panel of Figure 2). In this architecture, a hole (an explicit exception rule allowing specific traffic through the firewall) must be opened in the firewall to allow all users on the corporate LAN to request web pages from external web servers (the communications path labeled "1" in Figure 2A). In addition, a hole must be opened in the firewall to allow external web users on the Internet to request web pages from the web server on the corporate LAN (path "2" in Figure 2A).

With the services flowing through the Figure 2A example network as described above, we can determine what a motivated attacker would have to do to compromise critical assets on the corporate LAN. A significant vulnerability in the web server application running on the corporate web server may allow an attacker on the Internet to take over the server by sending malicious frames through the hole we had to open in the firewall (path “2” in Figure 2A). Because the now compromised web server is connected directly to the corporate LAN, attackers on the Internet can use it to target other workstations and servers on the corporate LAN without any restrictions. There are no firewalls between the web server and other assets on the corporate LAN, so any vulnerable service addressable from the compromised web server can be exploited by the attacker to take down additional targets. There is a single point of failure in the architecture in Figure 2A: a major vulnerability in the web services application can put the entire corporate LAN in jeopardy.

If properly configured, the DMZ architecture shown in Figure 2B (the right panel of Figure 2) greatly increases the security of our web services example. This time, all Internet-accessible servers are placed on the intermediate DMZ LAN segment and no assets on the corporate LAN are allowed to access the Internet directly. Instead of requesting web pages directly from servers on the Internet, user workstations send the requests to a web proxy server on the DMZ (path “1a” in Figure 2B). This requires a hole in “Firewall 2” (the firewall between the corporate LAN and the DMZ in Figure 2B) to allow the web services requests through to the proxy. When a request is made, the proxy retrieves the requested data from the Internet (path “1b” in Figure 2B) and then forwards it to the requesting user on the corporate LAN. In order for the proxy to retrieve web pages, a hole must be opened in “Firewall 1” (the perimeter firewall between the DMZ and the Internet in Figure 2B) to allow the web services requests from the proxy through to the Internet. In addition, a hole must also be opened in “Firewall 1” to allow web services requests to flow from the Internet to the DMZ. This firewall exception allows users on the Internet to retrieve web pages from the corporate web server on the DMZ.

With the services flowing through the more secure Figure 2B example network as described above, we can again determine what a motivated attacker would have to do to compromise critical assets on the corporate LAN. A significant vulnerability in the web server application running on the corporate web server may allow an attacker on the Internet to take over the server by sending malicious frames through the hole we had to open in “Firewall 1” (path “2” in Figure 2B). This time, however, the web server was placed on the DMZ, so the attacker does not automatically gain access to services on the corporate LAN through the compromised web server. In fact, a stateful firewall will block all unsolicited traffic from the DMZ to the corporate LAN because no ingress exceptions are required on the DMZ interface of “Firewall 2” to allow return traffic from the web proxy to the requesting user on the corporate LAN (path “1a” in Figure 2B)[2]. For this simplified example, there are no services accessible on the corporate LAN from the DMZ that would allow an attacker to compromise assets on the corporate LAN: the network compromise is effectively isolated to the DMZ.

[2] If a non-stateful firewall were used, an explicit ingress rule would have to be created to allow web services to flow from the DMZ to the corporate LAN. This hole in the firewall would allow, in addition to the return traffic associated with path “1a” in Figure 2B, unsolicited web services requests from the DMZ to the corporate LAN. Because of this, a non-stateful firewall would allow an attacker to probe for vulnerabilities in accessible web servers located on the corporate LAN.

Because servers and other assets on the DMZ are more exposed to electronic attack, the operating systems and firmware running on them should be as robust as possible. Security-related updates and patches should be applied to equipment located on the DMZ in a timely manner. These patches are usually provided by software or hardware vendors to fix identified vulnerabilities that may be exploited by attackers to gain control of critical assets. Patching accessible communications services removes these vulnerabilities and significantly improves the security of critical services.

We used a common IT problem to illustrate the benefits of the DMZ architecture, but the security benefits and improved attack isolation properties discussed above are extremely useful in a variety of network applications. Employing the DMZ architecture shown in Figure 1 is a valuable technique for improving the security of network interconnections. This is especially true whenever it becomes necessary to connect critical ICS networks to less regulated, risky network segments.

3.2. Install Cryptographic Security Modules to Block External Attacks from Insecure WAN Media

Any network link that exits the physically-secured perimeter may be vulnerable to electronic intrusion. Gaining physical or electronic access to a network link may allow an attacker to intercept data from, or write data to the networked assets serviced by the compromised link. The risk of an attack from the link infrastructure itself depends entirely on the physical security of the link media and/or the electronic security of the WAN infrastructure used to implement the link. In this document, we will refer to attacks originating from the WAN link infrastructure outside the protected physical perimeter as “external” attacks, and attacks originating from within the protected physical perimeter as “inside” attacks. Some examples of potentially externally-exploitable network link implementations follow:

- Internet connections are exposed to random or targeted attacks from anywhere in the world
- Dialup modem links are globally-accessible to any attacker with a functional telephone line
- Wireless network links are susceptible to data interception or injection from attackers sufficiently close to the antennas
- The security of leased line network infrastructures (e.g. public switched telephone network) is controlled by a third party and may not be sufficient for critical links
- Physically-accessible network media may be susceptible to wiretapping

Access control mechanisms implemented in the networked assets serviced by a communications link can be effective at blocking external attacks. For example, remote web or terminal services that use strong authentication and encrypt or properly obfuscate transmitted authentication credentials (e.g. passwords) are generally considered quite secure. Unfortunately, many of the communications protocols and user authentication schemes in common use on existing ICS networks are not sufficiently secure. For example, user login services often do not support password values that are complex enough to withstand password guessing attacks, and password values are often transmitted in a format that can be intercepted from a vulnerable WAN link and exploited by a potential attacker. Vulnerable WAN connections that provide access to insufficiently-secured communications services should be protected from external attacks with additional security technologies.

A very effective way to defend a vulnerable WAN link from external attacks is to install inline cryptographic security modules at every point of the link that enters/leaves a secured network perimeter. Figure 3 shows several vulnerable point-to-point WAN links protected by cryptographic security modules on both ends of each link (represented by lock and key icons in the figure).

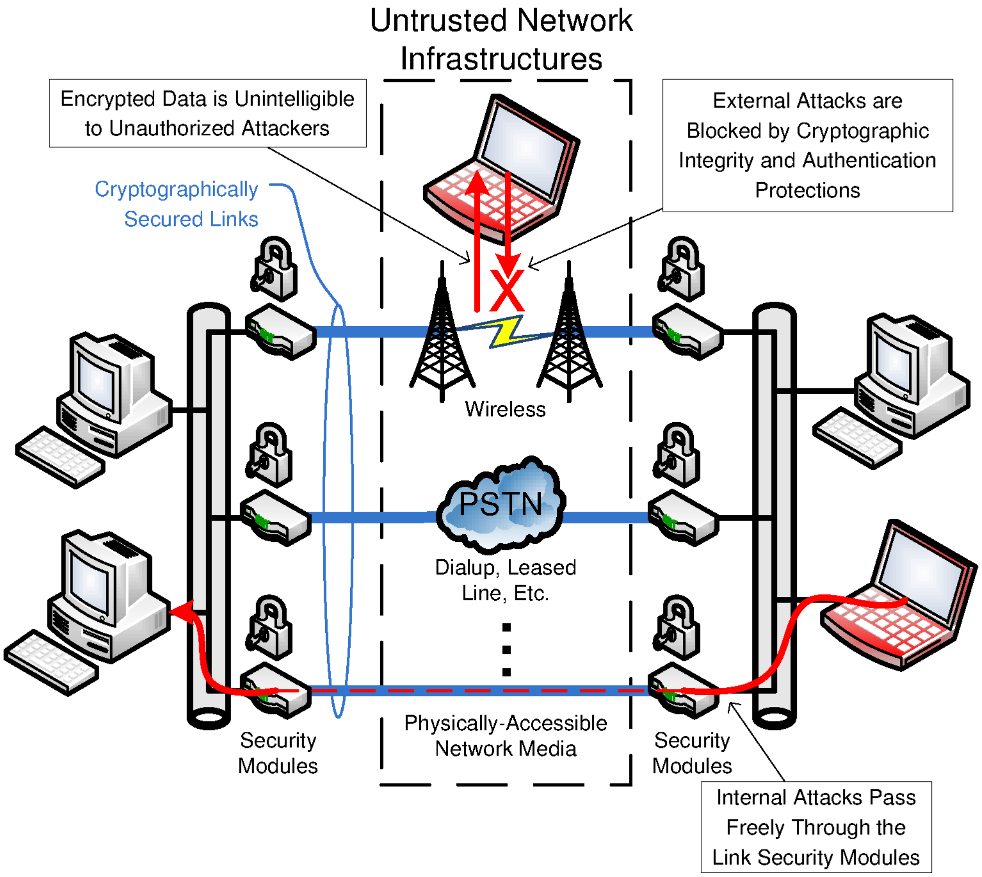

Untrusted Network Infrastructures
Encrypted Data is Unintelligible to Unauthorized Attackers
External Attacks are Blocked by Cryptographic Integrity and Authentication Protections
Cryptographically Secured Links
Wireless
PSTN
Dialup, Leased Line, Etc.
Physically-Accessible Network Media
Security Modules
Security Modules
Internal Attacks Pass Freely Through the Link Security Modules

Figure 3 Cryptographic Link Security

The security modules shown in Figure 3 encode data leaving the protected network perimeter with cryptographic security algorithms, and decode received data prior to allowing it into the protected perimeter. The cryptographic encoding/decoding protects the data in transit and prevents unauthorized data from entering the protected network perimeters. It is extremely important to point out that the cryptographic protections applied by the perimeter security modules very effectively block external attacks originating from the untrusted WAN infrastructure, but are generally not configured to protect against inside attacks from within the protected network perimeter. As shown in Figure 3, the perimeter security modules pass data between the protected network perimeters without authenticating or authorizing users within the trusted network segments. Because of this, it is critical to enforce the principle of least privilege within the trusted network infrastructure to reduce the risk of insider attacks and attack propagation amongst “trusted” network assets in the event of a security breach. Link security modules do not replace the need for implementing secure network planning and design, strong user authentication, effective network traffic filtering, network monitoring and other closed network security principles.

Modern cryptographic security modules provide a combination of the following three perimeter security functions:

- **Encryption:** ensures that data cannot be intercepted and interpreted by unauthorized entities
- **Integrity:** ensures that unauthorized data modifications are detected
- **Authentication:** ensures that data transmissions by unauthorized entities are detected

Encryption mechanisms should be implemented on all vulnerable WAN links that carry unprotected passwords or other sensitive data to prevent eavesdropping. An external attacker able to intercept data from the untrusted WAN infrastructure will not have knowledge of the binary key used to encrypt the data and, thus, will be unable to decrypt and interpret the intercepted messages (see Figure 3). Encryption can also make it difficult, but not impossible for an external attacker to successfully modify or inject data to negatively affect assets on the protected network segments. The decryption process will usually turn the injected attack frames into unintelligible data that will be ignored by networked assets. Though a successful injection or data modification attack against an encrypted link is less likely, potentially damaging attack data may still make it through your network perimeters. It is much more secure to employ dedicated cryptographic authentication mechanisms to keep unauthorized, malicious data off of protected networks. Cryptographic integrity and authentication[3] mechanisms do not provide data confidentiality but do prevent modified or injected data from breaching your network perimeter. As shown in Figure 3, all traffic transmitted or altered by the external attacker from within the untrusted WAN infrastructure is blocked at the network perimeter by cryptographic integrity and authentication protections.

There are a wide variety of cryptographic security modules on the market that can simultaneously integrate encryption, integrity, and authentication protections to lock down and secure a critical communications link. There are products available that are compatible with the most common communications links found in ICS networks. Most WAN link termination equipment (e.g. data service units, radios, or dialup modems) can accept either serial or TCP/IP data, so cryptographic modules compatible with serial or TCP/IP data streams will handle the majority of ICS link security needs.

You can cryptographically secure any TCP/IP traffic flowing over vulnerable network links by installing Internet Protocol Security (IPSec) based Virtual Private Network (VPN) devices at the ends of each link (e.g. at each point that the network link leaves the physically-protected network perimeter). IPSec-based VPNs can be configured to apply state-of-the-art encryption and/or authentication to protect data payloads leaving/entering the protected network perimeters.

Inline serial cryptographic modules are also available that provide, at minimum, strong encryption. Many serial links operate at relatively slow baud rates that may not be compatible with state-of-the-art message authentication techniques. In order to authenticate each message individually, a relatively large Message Authentication Code (MAC) must be appended to the end of each data frame (usually 8-16 bytes each). Slow serial lines may not accommodate this cryptographic overhead. Because of this, many serial

[3] Most cryptographic message authentication techniques also guarantee the integrity of the data. From this point forward, we will not differentiate between the two concepts: “authentication” will also imply integrity protections unless otherwise stated.

cryptographic security modules do not support strong message authentication and rely, instead, on less bandwidth intensive authentication mechanisms (e.g. periodic session authentication). If a vulnerable link can accommodate the accompanying data overhead, you should choose serial cryptographic modules that support strong message authentication. If not, you will have to choose slightly less secure modules that minimize cryptographic overhead. Finally, you must also ensure that the serial cryptographic modules that you choose can support the baud rate of the network link that you wish to secure: many serial security modules have a relatively low maximum baud rate (e.g. 19,200 or 38,400 bits/second).

3.3. Monitor Network Activity on Vulnerable Links

Standalone intrusion detection systems (IDS) are widely-available for IP networks. These usually take the form of a standalone PC running a dedicated IDS software package like Snort. PC-based IDS equipment is extremely inexpensive: adequately-powered PCs are cheap and effective open source IDS software packages like Snort are free. The IDS is usually placed at critical network choke points through which attack traffic is likely to flow. An IDS can be attached to a network monitoring port on a router or intelligent switch from which it can analyze all traffic flowing through the router/switch. A modern IDS can be configured to automatically detect common attack signatures like port scans, known vulnerability exploit traffic (e.g. buffer overflow attacks or worm propagation traffic), or abnormal network traffic (e.g. the corporate web server is trying to access file shares on the corporate network). When a potential attack is observed, the IDS will automatically log the event and notify security personnel.

IP Routers and firewalls also produce extensive logging that can be consolidated in a central location (e.g. in a central log repository via the syslog protocol). These devices can often be configured to log a time-stamped record of each TCP/IP session routed through the device or filtered/dropped by the device. These activity logs can be analyzed by security personnel manually or by using a Security Information and Event Management (SIEM) platform to detect suspicious network activity in areas on the network where an automated IDS is not present.

Monitoring network activity on a serial network segment can be much more difficult. Universal, dedicated IDS technologies simply do not exist for non-TCP/IP, serial networks. This is a direct consequence of the fact that there is no unifying communications protocol carried on most serial networks: the network may be carrying DNP3 SCADA protocol frames, ASCII terminal data, or any number of serial protocols. It is, however, usually possible to monitor user activity directly in serially-connected devices. Border serial communications concentrators like protocol gateways, modems, SCADA data concentrators, or serial engineering access port switches form serial communications choke points and are convenient locations to analyze activity logs. The visibility offered by common serial ICS networking devices varies greatly, but it is usually possible to observe useful events like connection logs (e.g. from a modem), engineering access activity (e.g. successful or failed user logins), or communications statistics (e.g. reception of malformed protocol frames).

Vulnerable communications choke points that carry network traffic that can be used to negatively affect the reliable operation of critical control system functions should be routinely monitored whenever possible. SIEM products should be considered to centralize the monitoring and analysis of distributed logging and event data.

4. Specific ICS Security Recommendations

4.1. Secure or Remove Connections Between the Internet and ICS Network Segments

Connections between the Internet and critical ICS network segments create extremely attractive remote electronic attack points. Vulnerabilities on these connections can create globally-accessible avenues through which attackers can, often anonymously, target critical ICS assets. It is often very easy to acquire the public IP addresses assigned to a potential target organization by querying the global Internet registry. It is therefore, prudent to assume that any diligent attacker has the information required to target these publicly-accessible connection points. Any connection, direct or indirect, between the Internet and critical ICS assets should be considered high risk. Such connection points should, ideally be removed entirely or combined into a single, highly-secured and regulated link as explained below.

The most common ICS network connection to the Internet occurs when corporate LAN segments are linked to critical ICS Supervisory Control and Data Acquisition (SCADA) networks. These connections serve as a convenient way to access current and historical ICS status data from the corporate LAN for purposes of regional energy management, billing, planning, asset management, and many other important corporate functions. Critical SCADA assets become indirectly exposed to Internet-born attacks when these corporate LAN segments are also connected to the Internet to provide corporate users access to web-based services like e-mail and web browsing. Attackers may be able to exploit enough vulnerabilities on the corporate network to eventually target critical SCADA assets.

If such connections must exist to implement required functionality and data sharing, then they must be considered high risk and secured and monitored accordingly. The DMZ architecture introduced in Figure 1 should be used to isolate the ICS SCADA network from the high risk Internet-connected corporate LAN. Figure 4 shows an example ICS network with an intermediate DMZ subnetwork (labeled ‘ICS DMZ’ in the figure) between an Internet-connected corporate network and the critical ICS network assets.

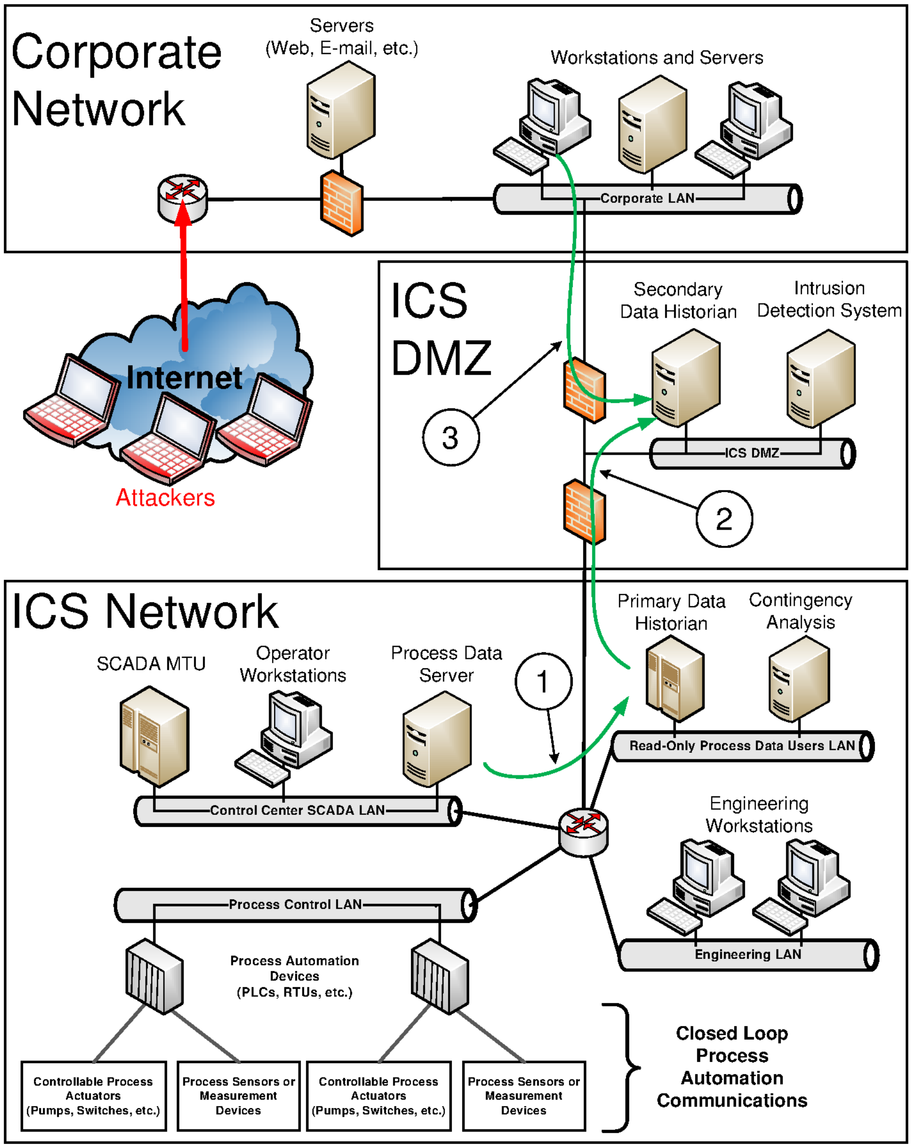

Corporate Network
Servers (Web, E-mail, etc.)
Workstations and Servers
Corporate LAN
Internet
Attackers
ICS DMZ
Secondary Data Historian
Intrusion Detection System
ICS DMZ
3
2
ICS Network
SCADA MTU
Operator Workstations
Process Data Server
1
Primary Data Historian
Contingency Analysis
Read-Only Process Data Users LAN
Control Center SCADA LAN
Engineering Workstations
Process Control LAN
Engineering LAN
Process Automation Devices (PLCs, RTUs, etc.)
Controllable Process Actuators (Pumps, Switches, etc.)
Process Sensors or Measurement Devices
Controllable Process Actuators (Pumps, Switches, etc.)
Process Sensors or Measurement Devices
Closed Loop Process Automation Communications

Figure 4 Example Network Architecture Employing a Demilitarized Zone (DMZ) to Further Isolate Critical ICS Network Segments from the Internet-Connected Corporate LAN

In Figure 4, the ICS network is segmented to separate users and devices with different roles, privileges, criticality, and communications requirements. Separate LAN segments have been created to isolate the control center SCADA assets, process automation assets, read-only process data users, and ICS device configuration/engineering workstations. Access control lists (ACLs) in the routers that connect the ICS network segments to one another ensure that the principle of least privilege is enforced by blocking non-vital communications services from flowing between network segments. Enforcing the principle of least privilege within the critical ICS network prevents an attack from propagating unhindered throughout critical segments of the ICS network in the event of a breach of the DMZ separating the ICS network from the corporate LAN.

The primary purpose of the ICS DMZ in Figure 4 is to securely pass process status data from the ICS network to the corporate network. In most ICS installations, an accurate archive of current and past events (e.g. control commands, alarms, etc.) and system status is of vital importance. Because of this, the primary data historian server (labeled ‘Primary Data Historian’ in Figure 4) performing this important data archival task is located on the ICS network rather than exposing it to a higher risk of attack on the ICS DMZ. To provide users on the corporate LAN access to the process data, an up-to-date, secondary copy of the data archive is maintained on a server on the ICS DMZ (labeled ‘Secondary Data Historian in Figure 4). The flow of process data is as follows:

1. Process state changes, events, and alarms are pushed from the process data server on the control center SCADA LAN to the primary data historian (the communications path labeled ‘1’ in Figure 4).
2. The primary data historian periodically pushes process data archive updates to a secondary data historian located on the ICS DMZ (path ‘2’ in Figure 4) so that the copy of the process data image remains sufficiently up-to-date.
3. Users on the corporate LAN request process data from the secondary data historian located on the ICS DMZ (path ‘3’ in Figure 4).

Ideally, the flow of data from the control center SCADA LAN to the read-only process data users LAN should consist of read-only updates (e.g. event-driven or periodic) from the process data server to the primary data historian. This prevents the need to provide access from the read-only process data users LAN to potentially-exploitable data retrieval services on the control center SCADA LAN. It also prevents access to commanded control points (e.g. breaker operation or pump control) from the read-only process data users LAN. For similar reasons, the flow of data from the read-only process data users LAN to the ICS DMZ should be read-only. Pushing data from the primary data historian to the secondary data historian eliminates the need to open a hole in the firewall that allows unsolicited data requests from the ICS DMZ to the critical ICS network.

In Figure 4, users or devices performing critical ICS functions that require read-only access to the process data, but do not require process control privilege (e.g. contingency analysis personnel) are isolated on the read-only process data users LAN. Users and devices on this network segment request process data directly from the primary data historian. If there is no extensive need for critical, read-only functions on the ICS network, then the read-only process data users LAN can be eliminated and the primary data historian can be placed on the control center SCADA LAN. This arrangement also eliminates the need for a separate process data server on the control center SCADA LAN: process data flows directly from the SCADA master terminal unit (MTU) to the primary data historian, which pushes data up to the secondary data historian on the ICS DMZ.

In Figure 4, an IDS has been placed on the ICS DMZ. The IDS is critical for detecting suspicious network activity that may be the precursor to an electronic attack that may eventually breach the DMZ and put critical ICS assets in jeopardy.

Firewall exception rules that permit services other than the read-only transfer of SCADA data from the ICS network to the corporate network, as described above, should be carefully considered prior to implementation. Each additional service running on the DMZ or ICS network that is accessible from the corporate network creates another potential point through which electronic attacks may flow into the protected ICS network. Services like Internet-connected energy management systems and third party remote engineering/monitoring that require access to assets on the DMZ, or assets on the protected ICS network must be cryptographically authenticated prior to granting access to protected network segments.

The list provided below contains some important recommendations that should be applied to improve the security of the DMZ architecture shown in Figure 4.

- Stateful firewalls should be used to isolate the ICS DMZ. This is especially important for the DMZ connection to the ICS network.
- Use different protocols to transfer data and services from the ICS network to the DMZ, and from the DMZ to the corporate network to protect against single vector attacks that may breach the DMZ (e.g. punch through the DMZ and directly connect the corporate LAN and the ICS network).
- Inbound and outbound firewall exception rules should enforce the principle of least privilege. Only essential services should be allowed through the firewalls and all permitted services should be locked down to the specific IP addresses and TCP ports (source and destination) required to implement the intended functionality.
- The transfer of external files from the corporate LAN to the ICS network should be minimized or eliminated. E-mail, external web access, and file sharing/transfer can be dangerous sources of malware (e.g. viruses, worms, or Trojan horse applications).
- All Internet-based communications flowing into the ICS network must flow through the ICS DMZ. Concentrating these risky connections at a single, highly-defended point makes it easier for security personnel to secure and monitor them. Additional connections between the Internet and critical ICS network assets (e.g. for implementing direct SCADA or engineering access WAN connections) should be eliminated.
- All connections between the corporate network and ICS network must flow through the ICS DMZ. Dual-homed PCs that bridge the two networks and bypass the security of the DMZ architecture must not be allowed.

4.2. Secure or Remove Vulnerable Dialup Modem Links

Despite their relatively dated technology, dialup modems communicating over public switched telephone network (PSTN) lines are still in wide use in ICS networks. Dialup links are an economical way to implement intermittent remote access functions like primary or backup engineering access, third party maintenance, or remote file transfers (e.g. retrieving system event recordings). Dialup modem use is especially prevalent in distributed ICS infrastructures like those that monitor and control the electric power grid.

Dialup modem links often provide direct access into the heart of an ICS network by connecting directly to one or more critical ICS assets. For example, an engineering and monitoring PC in a remote electric power system substation can be made remotely-accessible by connecting a dialup modem to one of its serial ports and setting up serial TCP/IP access on the port using the point-to-point protocol (PPP). Remote PC access services like Microsoft’s Remote Desktop or Symantec’s PCAnywhere can be configured to allow the remote user to perform virtually any of the duties they could perform while sitting in front of the physical computer. Dialup-accessible services often include the ability to reprogram critical ICS equipment or operate critical ICS control points (e.g. pumps, switches, or electric power system breakers).

Dialup modem links utilizing the PSTN infrastructure are globally-accessible. Furthermore, the critical services and capabilities that these remote dialup links provide access to are often not sufficiently secured. Weak or non-existent user authentication mechanisms often make it fairly easy for a motivated attacker to maliciously affect the safe and reliable operation of critical control system assets via the remote dialup connection. Vulnerable remote dialup links that allow modification of critical ICS functions should be removed or secured with additional cryptographic security protections. The following list provides some examples of acceptable methods for adding cryptographic security to dialup modem links:

- Replace insecure modems with models that include strong cryptographic authentication mechanisms (e.g. via a keyed challenge/response session initiation dialog or application of Hash-based Message Authentication Codes (HMAC) to transmitted data frames) .
- Install inline, serial cryptographic security modules to secure existing modems. There are models available that are compatible with common ICS communications protocols and include both cryptographic authentication and encryption mechanisms.

It is also possible to use serial to Ethernet converters in conjunction with Internet Protocol Security (IPSec) Virtual Private Networking (VPN) products to create cryptographically-secured serial dialup access points. This methodology is not recommended because such connections create a globally-accessible access point to all equipment attached to the connected network segment. Even though the link is secured with strong cryptographic protections, “backdoor” remote network access points can be very difficult to regulate and monitor.

4.3. Secure or Remove Vulnerable Wireless Radio Links and Network Access Points

Wireless communications products are quite common in many ICS applications. For example, point-to-point, microwave or spread spectrum radio links are an economical way to connect isolated ICS equipment to wired telecommunications access points, or directly to a SCADA control center. Examples include connecting isolated electric power system substations or ICS monitoring and control equipment on the plant floor/yard. It is also becoming more common to use 802.11x based network access points in ICS environments. Wireless network access is a convenient tool for connecting mobile personnel to critical ICS services. Examples include wireless order processing or factory inventory systems.

Wireless antennas radiate transmitted signals to locations outside the secured physical perimeter. These signals can be demodulated with compatible radio equipment, putting all data transmitted over wireless links at risk of interception. In addition, wireless antennas are exposed to reception of malicious signals transmitted from outside the secured physical perimeter. Malicious transmissions can disable the link (e.g. by jamming the signal) or send malicious attack data to connected equipment. It is possible to protect these links from unauthorized outside intrusion (e.g. transmission of malicious data to connected equipment), but it is not possible to eliminate the possibility of intended or accidental external jamming of wireless reception. Backup communications or response procedures should be in place to mitigate the consequences of the loss of any wireless communications link: such losses should not cause an appreciable reduction in safety or reliability of the controlled process.

Any wireless link that carries vulnerable communications services that may be exploited by an outside attacker to affect the safe, reliable operation of critical control system functionality should be secured with strong authentication mechanisms or replaced with a more secure, wired link. SCADA or engineering access links that allow commanded control of, or reprogramming of critical ICS equipment are of particular concern: SCADA protocols do not typically include security mechanisms to prevent unauthorized command injection and user authentication mechanisms common in ICS device engineering services are often insufficient to prevent unauthorized alteration of critical device settings. Cryptographic link security modules should be employed on vulnerable wireless links (see Figure 3) to lock out unauthorized intrusion. Cryptographic security is sometimes available directly in the radio, eliminating the need for additional inline cryptographic security modules. The employed security solution should include strong cryptographic session and/or data authentication mechanisms, and should include encryption functions whenever sensitive data and/or passwords are transmitted in-the-clear (unencrypted and vulnerable to interception) over the wireless link.

802.11x-based wireless access points that provide direct or indirect access to critical ICS network segments should be isolated from the protected, wired ICS network with a dedicated wireless access DMZ. Servers and proxies on the DMZ provide the required ICS functions that must be made accessible via the wireless access point (e.g. order entry or inventory services). The DMZ should stand between the wireless access point and the protected network segments as shown in Figure 5. Note that we used a single, three-interface firewall for simplicity.

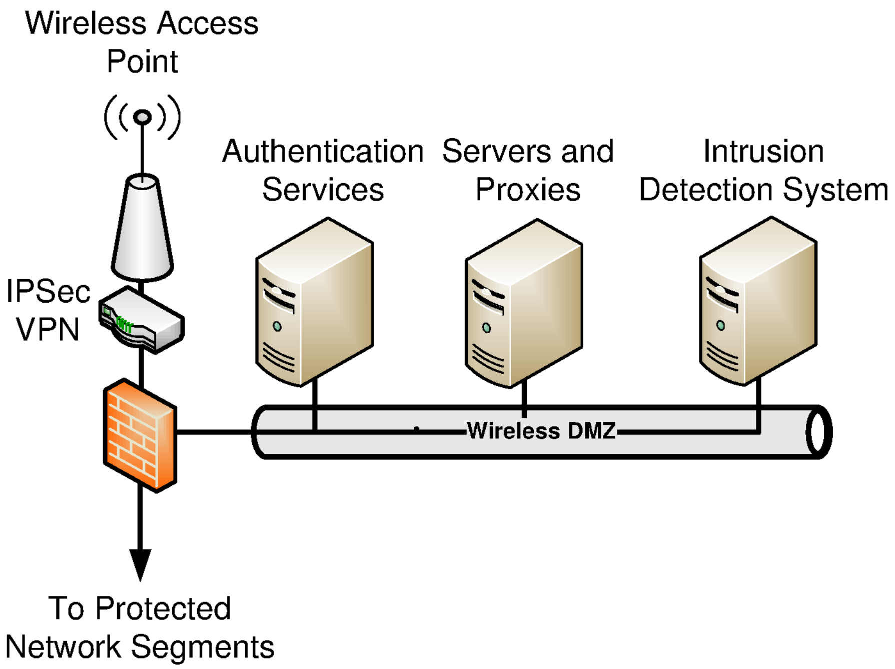


Figure 5 a DMZ Should be Used to Isolate 802.11x Wireless Access Points from Critical ICS Network Segments.

In Figure 5, we have placed an IDS on the Wireless DMZ to detect suspicious network activity and to warn of potential electronic attacks. IDS equipment is inexpensive and should be placed on all IP network chokepoints that connect to critical ICS network segments and have an appreciable exposure to outside attack.

The wireless access point shown in Figure 5 should be the only 802.11x-enabled device allowed on the wired ICS network: any device with an enabled 802.11x wireless interface should be physically disconnected from the protected, wired ICS network. This policy prevents vulnerable wirelessly-accessible attack points behind the network perimeter defenses (e.g. due to a PC with a wireless interface card in ad-hoc mode connected to a critical TCP/IP network segment). For ICS network environments, it is usually safest to designate a device is either permanently wireless or permanently wired. Permanently wireless devices may have enabled 802.11x wireless interface cards/adaptors but are never allowed to connect to the wired ICS network. If necessary, you can remove or disable all physical/electrical TCP/IP interfaces (e.g. Ethernet ports) on the device to enforce this policy. Permanently wired devices, on the other hand, may directly connect to wired ICS TCP/IP network segments (e.g. via a physical Ethernet port) but should never contain 802.11x wireless interface cards/adaptors.

All wireless connections must be authenticated with proven technologies prior to being allowed access to services and assets on a wired network; the WEP, WPA, and/or WPA2 security provisions built into the 802.11x standards are not sufficient. In Figure 5, the server labeled ‘Authentication Services’ authenticates incoming wireless access requests (e.g. using RADIUS or domain services). An IPSec-based VPN device placed between the access point and the DMZ as shown in Figure 5 cryptographically secures all traffic entering/leaving the wireless access point. Some wireless access point products include IPSec services, thus combining the separate ‘Wireless Access Point’ and ‘IPSec VPN’ devices shown in Figure 5.

4.4. Prevent Access to Unprotected Physical Network Media

Physically-accessible, wired network media links that leave the secured network perimeter may be susceptible to outside intrusion via wiretapping. Convenient media access points like communications cabinets and unused communications ports should be placed inside the physical security perimeter (e.g. in a telecommunications room in a locked building or inside a protected fence line) to reduce the chance of outside access to critical networks. If possible, wired network media runs should be buried or placed out of convenient reach (e.g. attached to overhead poles or electric power transmission/distribution structures).

It is almost never possible to completely eliminate the chance of physical access to network media. This is especially true for wide area network link implementations. Oftentimes the most convenient and cost effective way to reduce the chance of electronic intrusion through a physically-exposed network link is to cryptographically secure the data carried by the link. Inline cryptographic security modules can be used on overly-exposed, critical network links to defeat data interception and malicious injection attacks through media splices or wiretaps (see Figure 3).

5. Conclusion

Electronic attacks are most likely to enter protected networks through communications links that are significantly exposed to unauthorized outsiders. Electronic access points that create direct or indirect connections between critical ICS network segments and publicly-accessible network infrastructures like the Internet or PSTN are of particular concern. In this publication, we presented network architectures and cryptographic security technologies that can be employed to block unauthorized network traffic from entering protected ICS network segments through these high-risk network connections.